Deep Learning Atmospheric Features

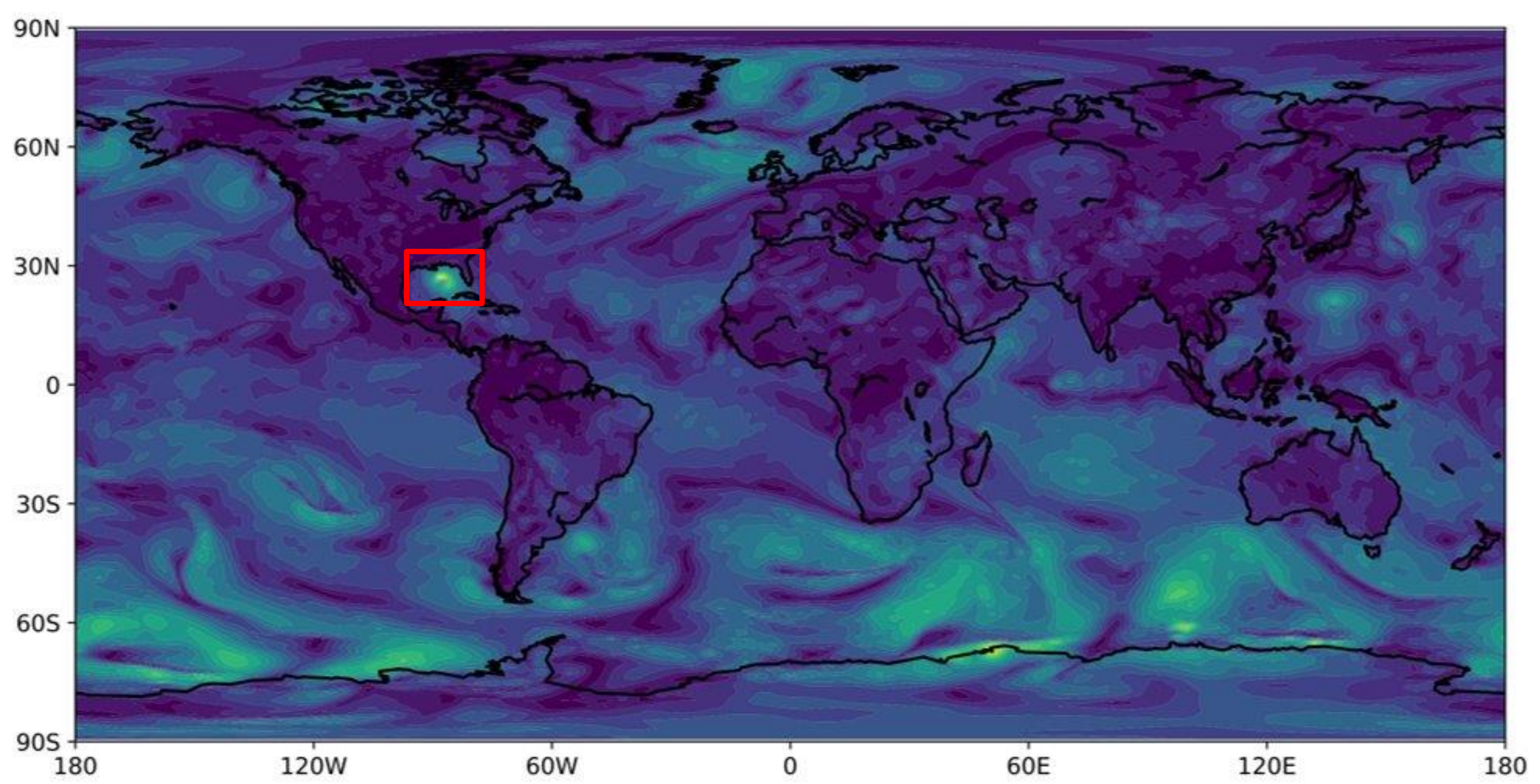
Daniel Galea
d.galea@pgr.reading.ac.uk
PhD @ University of Reading

The Problem

- Weather and Climate models currently produce TBs of data, and that will continue to increase as they get better at finer resolutions
- We need an automatic way to filter the data to only store and analyze the important bits of data
- First such important bit of data is a Tropical Cyclone (i.e. a Hurricane)

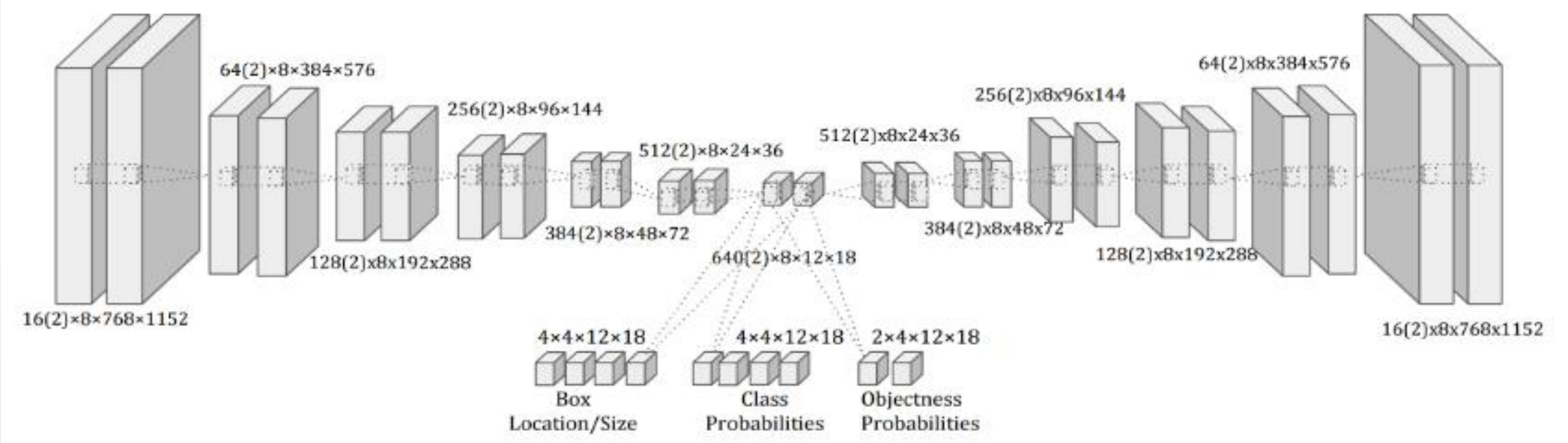
The Problem

- "*A tropical cyclone in which the maximum sustained surface wind (using the U.S. 1-minute average) is 64 kt (74 mph or 119 km/hr) or more.*" - NOAA NHC
- Wind Speed Field [Hurricane Katrina (2005-08-28)]:



Previous Work

- Racah et al. (2016) produced an autoencoder model to replicate the input fields and perform detection and localization at the bottleneck



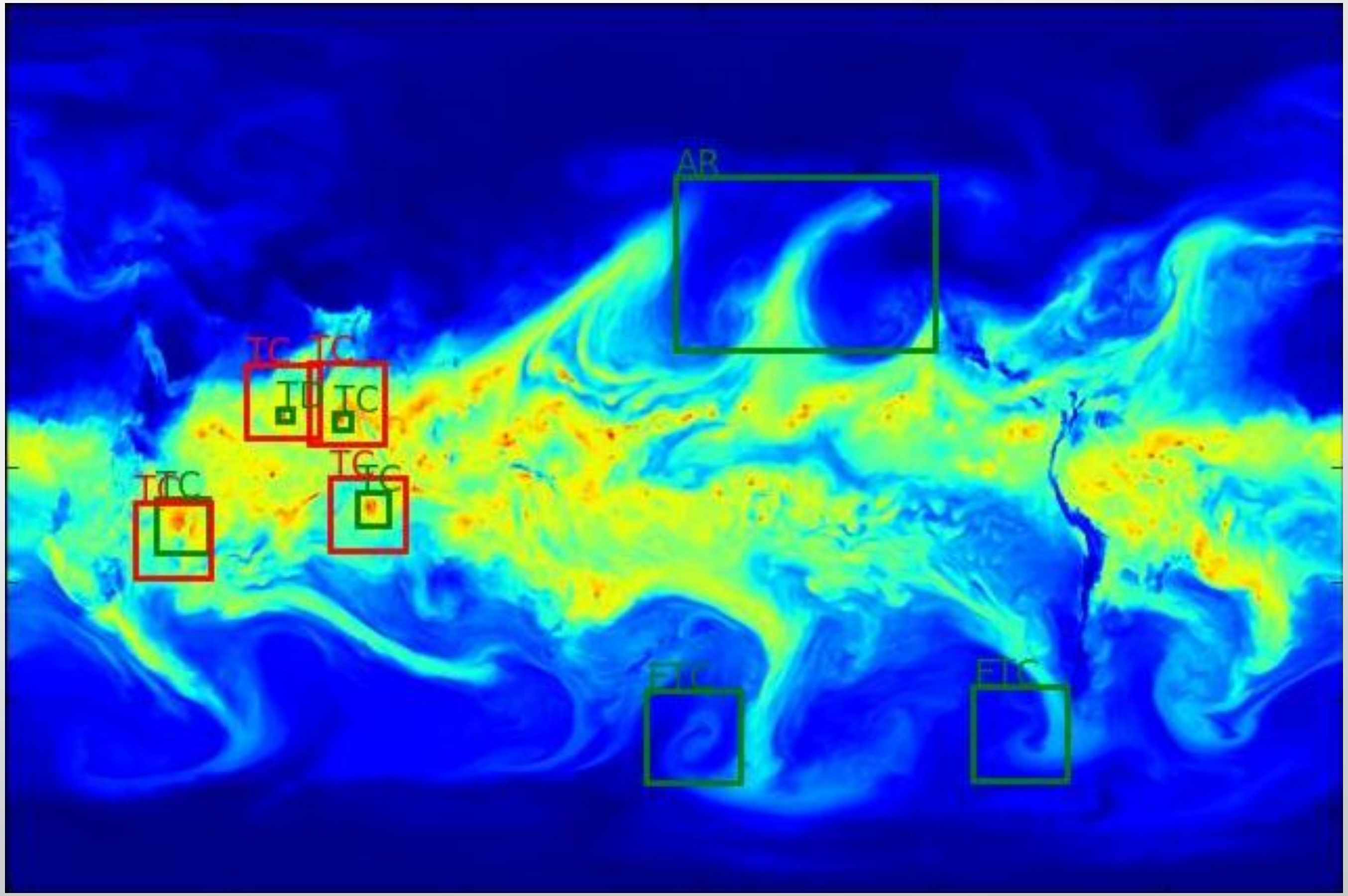
- Variables Used: PRECT (Precipitation Rate); PS (Surface Pressure); PSL (Sea Level Pressure); QREFHT (Reference Height Humidity); TS, T200, T500, TREFHT (Temperatures); TMQ (Total water vapor); UBOT, U850, VBOT, V850 (Winds); Z1000, Z200, ZBOT (Geopotential)

Racah, E., Beckham, C., Maharaj, T., Kahou, S. E., Prabhat, & Pal, C. (2016). ExtremeWeather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events, (Nips), 1–12. Retrieved from http://arxiv.org/abs/1612.02095

Previous Work

- Ground truth labels created using TECA: Toolkit for Extreme Climate Analysis. Some issues with false positives.

	Intersection Over Union	
	0.1	0.5
Accuracy	24.74%	15.53%



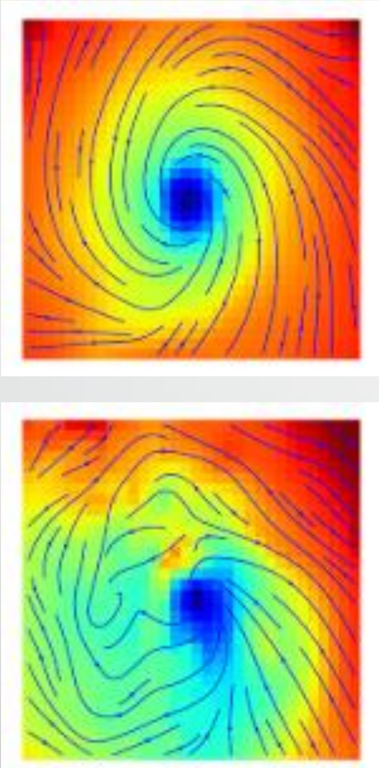
Racach, E., Beckham, C., Maharaj, T., Kahou, S. E., Prabhat, & Pal, C. (2016). ExtremeWeather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events, (Nips), 1–12. Retrieved from http://arxiv.org/abs/1612.02095

Previous Work

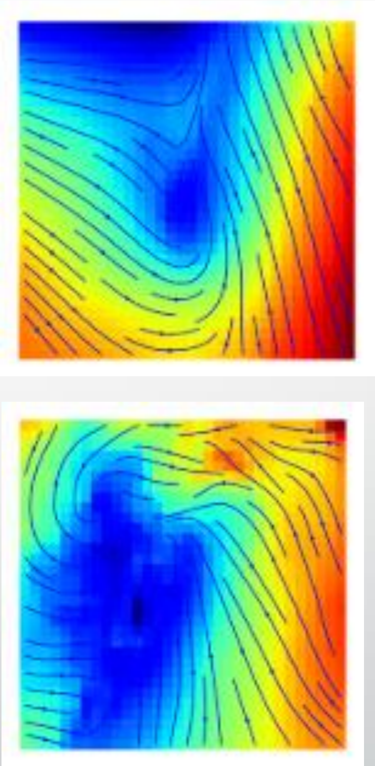
- Liu et al. (2016) uses a deep learning network to classify cropped images of TCs or non-TCs.
- Ground truth labels created using TECA: Toolkit for Extreme Climate Analysis. Some issues with false positives.
- Variables Used: PSL (Sea Level Pressure); T200, T500 (Temperatures); TMQ (Total water vapor); UBOT, U850, VBOT, V850 (Winds)
- Architecture used:
  - Conv2D: 5x5 @ 8; MaxPool: 2x2
  - Conv2D: 2x2 @ 16; MaxPool: 2x2
  - Dense: 50
  - Dense: 2

Previous Work

- Results of 99% accuracy were obtained.



Correctly Classified TC



Incorrectly Classified TC

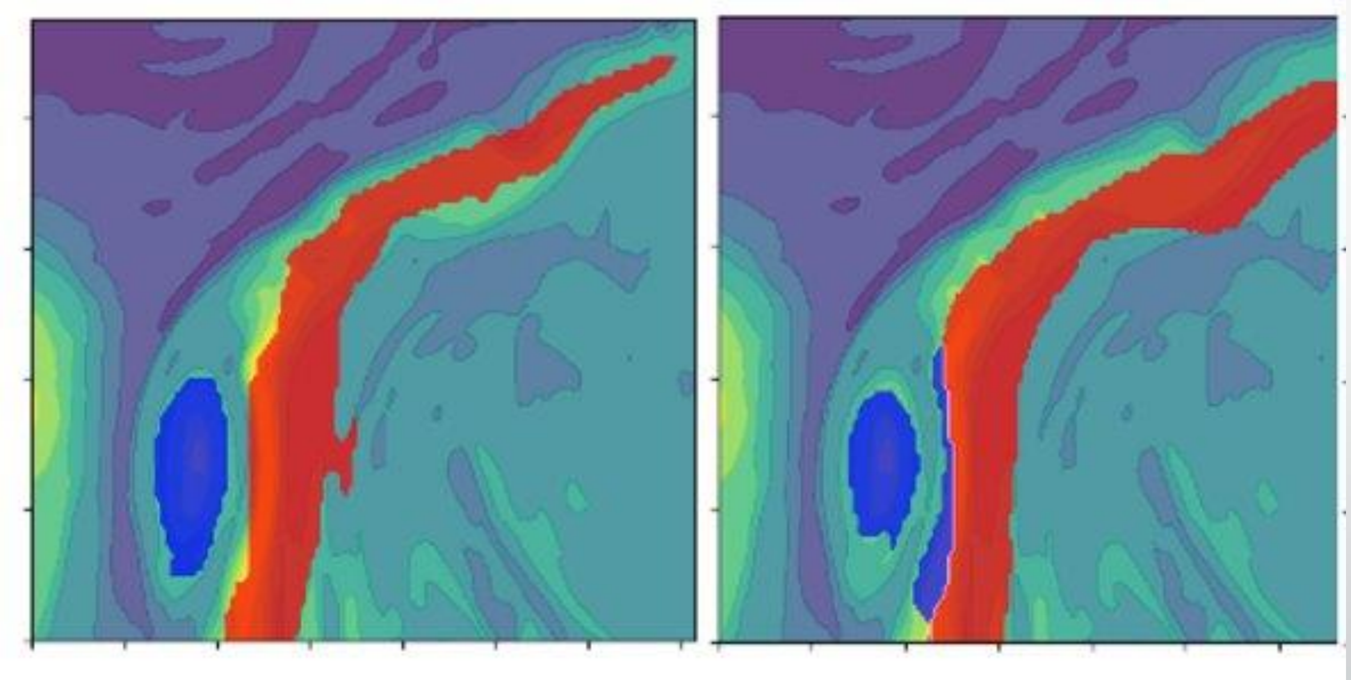
Liu, Y., Racah, E., Prabhat, Correa, J., Khosrowshahi, A., Lavers, D., ... Collins, W. (2016). Application of Deep Convolutional Neural Networks for Detecting Extreme Weather in Climate Datasets. https://doi.org/10.475/123

Previous Work

- Mudigonda et al. (2017) uses a deep learning network to perform image segmentation (by pixels)
- Use images with >10% non-background pixels.
- Ground truth labels created using TECA: Toolkit for Extreme Climate Analysis. Some issues with false positives.
- Variable used: IWV (Integrated Water Vapour)
- Architecture: DenseNet

Previous Work

- 92% accuracy was obtained.

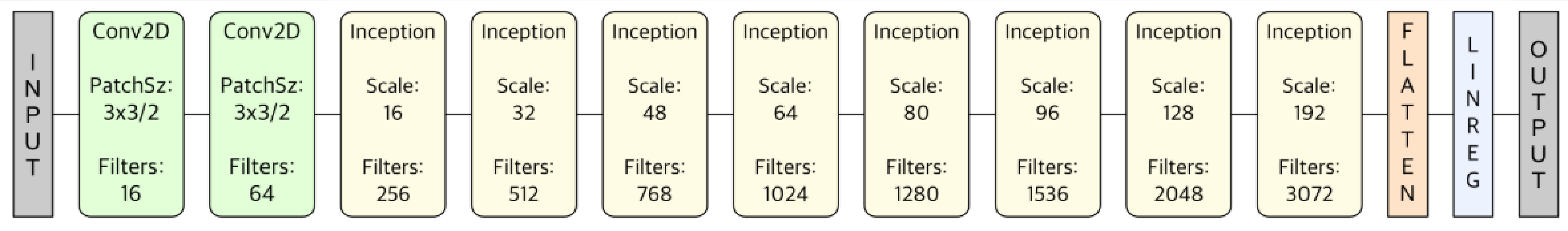


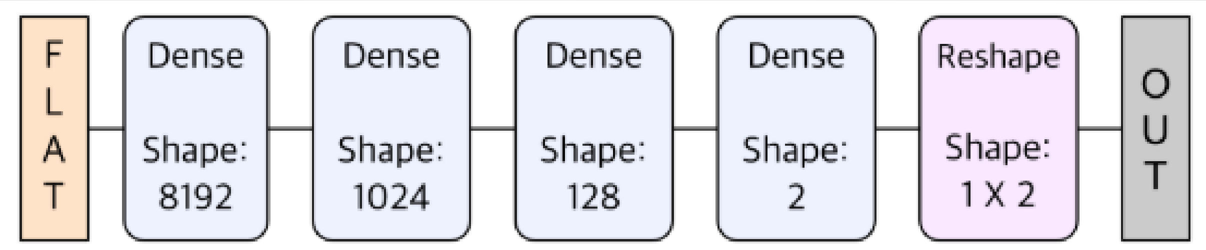
(a) Ground Truth (b) Prediction

Mudigonda, M., Kim, S., Mahesh, A., Kahou, S., Kashinath, K., Williams, D., … Prabhat, M. (2017). Segmenting and Tracking Extreme Climate Events using Neural Networks. 31st Conference on Neural Information Processing System, (Deep Learning for Physical Climate), 1–5. Retrieved from https://dl4physicalsciences.github.io/files/nips_dlps_2017_20.pdf

Previous Work

- Hong et al. (2017) uses a deep learning network to locate the centre of the storm (the Eye) from four IR images from the Northwest Pacific basin.
- Ground truth positions found using Japan Meteorological Agency's dataset.
- RMSE of 0.02 (~74km) was obtained.
- Architecture:





Mudigonda, M., Kim, S., Mahesh, A., Kahou, S., Kashinath, K., Williams, D., … Prabhat, M. (2017). Segmenting and Tracking Extreme Climate Events using Neural Networks. 31st Conference on Neural Information Processing System, (Deep Learning for Physical Climate), 1–5. Retrieved from https://dl4physicalsciences.github.io/files/nips_dlps_2017_20.pdf

Current Work

- Previous work has all been done on some sort of cropped image, usually with the Tropical Cyclone in the centre of the image
- If a model's output is cropped into such images and each image is tested, a large number of computations will be required, thus slowing the climate model down.
- Would like to find a model that performs detection of Tropical Cyclones to as close to a full global image as possible to keep running costs down

Current Work

- Obtained data from ECMWF's ERA-interim dataset (1979-2016)
- Using Kevin Hodges (UoR) TRACK algorithm to locate Tropical Cyclones
- Using MSLP and wind speed fields in the dataset, which is split 50:50 between cases with a TC present and not.
- Only taking into account the Tropics
- Current architecture:
  - Conv2D: 5x5 @ 16; MaxPool: 2x2
  - Conv2D: 2x2 @ 32 (L2 weight regularization); MaxPool: 2x2
  - Conv2D: 2x2 @ 64; MaxPool: 2x2
  - Flatten
  - Dense: 128
  - Dense: 64 (L2 weight regularization)
  - Dense: 32

Current Accuracy: ~69% classifcation

Current Work

- Are the Tropical Cyclones to small to pick out in a global picture?
- Have split a TC track into six parts to mirror lifecycle of Tropical Cyclone and infer strength of Cyclone at the time
- Will evaluate model on each of the parts of the track and obtain diagnostics for each part to check if one part of the lifecycle is better picked out
- If so, will try to optimize model to parts that do not perform well or develop models for each part
- Collaboration with Nvidia/Oracle to test model on GPUs
- Other outstanding issue:
    - Using Horovod/Dask and Singularity containers in tandem for when size of data used increases